Di Wang

Al Khawarizmi Building 1, Room 4341 King Abdullah University of Science and Technology Thuwal Kingdom of Saudi Arabia, 23955-6900

Email: di.wang@kaust.edu.sa https://shao3wangdi.github.io/ Tel: +966 (012) 8080645

WORK EXPERIENCE

King Abdullah University of Science and Technology Computer Science Division of CEMSE Assistant Professor Director of Privacy-Awareness Responsibility and Trustworthy Lab

Thuwal, Saudi Arabia January 2021-Current

August 2020

EDUCATION

State University of New York at Buffalo Buffalo, NY, United States Ph.D. in Computer Science and Engineering

Advisor: Dr. Jinhui Xu

Dissertation: Some Fundamental Machine Learning Problems

in the Differential Privacy Model

Western University (University of Western Ontario) London, ON, Canada M.S. in Mathematics October 2015

Jinan, Shandong, China Shandong University B.S. in Mathematics and Applied Mathematics June 2014

CURRENT STATUS

Citizen of China

RESEARCH INTERESTS

Private Data Analytics

Differential privacy, privacy-preserving machine learning, privacy-preserving data mining, privacy attack

Trustworthy Machine Learning

Robust estimation, fairness in machine learning, machine unlearning interpretable machine learning, causality

Statistical Learning Theory

Large scale optimization, high dimensional optimization, statistical estimation, learning theory, quantum machine learning

Biomedicine and Healthcare

Trustworthy issues in digital healthcare, biomedical imaging and bioinformatics

RESEARCH EXPERIENCE

University of California at Berkeley
Simons Institute for the Theory of Computing
Spring 2019

Data Privacy: Foundations and Applications

Visiting Graduate Student

Harvard University Cambridge, MA Harvard University Privacy Tools Project June to August 2018

Research Graduate

Boston University Boston, MA
Visiting Student June to August 2018

Mentor: Dr. Adam D. Smith

State University of New York at Buffalo Buffalo, NY Research Assistant August 2015 to December 2020

Supervisor: Dr. Jinhui Xu

HONORS and **AWARDS**

• CSE Best Doctoral Dissertation Award in 2020, SUNY at Buffalo.

- SEAS Dean's Graduate Achievement Award in 2019, SUNY at Buffalo.
- Best CSE Graduate Research Award in 2018, SUNY at Buffalo.
- ICML Travel Award, 2019.
- NeurIPS/NIPS Travel Award, 2019, 2018, 2017.
- Western Graduate Research Scholarship, Western University, 2014-2015.
- Algebraic Geometry Summer School Scholarship, ENCU, Shanghai, 2013.

TEACHING EXPERIENCE

- Instructor. CS3xx: Contemporary Topics on Cybersecurity-Differential Privacy. Fall 2021, KAUST
- Instructor. Short Course: Selected Topics in Differentially Private Machine Learning and Statistics, 5th-7th January 2021, School of Computer Science and Technology, East China Normal University.
- Instructor. CSE574/474: Introduction to Machine Learning, Summer 2019, State University of New York at Buffalo.
- **Teaching Assistant.** CSE574/474: Introduction to Machine Learning, Spring 2018, State University of New York at Buffalo.
- Teaching Assistant. CSE531/431: Analysis of Algorithm, Fall 2017, Spring 2017, Fall 2016, Spring 2016, State University of New York at Buffalo.
- **Teaching Assistant.** CSE115: Introduction to Computer Science for Majors I, Fall 2015, State University of New York at Buffalo.
- **Teaching Assistant.** MATH 1229A: Methods of Matrix Algebra, Summer 2015, Winter 2015, Western University.
- Teaching Assistant. MATH 1225B: Methods of Calculus, Fall 2014, Western University.

Fundings

• Sole PI, 1,600,000 USD, KAUST Baseline, 2021-2026

SELECTED PUBLICATIONS

- 1. **Di Wang** and Jinhui Xu. On Sparse Linear Regression in the Local Differential Privacy Model. *IEEE Transactions on Information Theory*, Volume 67, No. 2, Pages 1182-1200, Feb. 2021.
- Di Wang, Marco Gaboardi, Adam Smith and Jinhui Xu. Empirical Risk Minimization in the Noninteractive Local Model of Differential Privacy. *Journal of Machine Learning Research*, Volume 21, 200 (2020), Pages 1-39.
- 3. Di Wang*, Xiangyu Guo*, Shi Li and Jinhui Xu. Robust High Dimensional Expectation Maximization Algorithm via Trimmed Hard Thresholding. *Machine Learning* 109, 2283-2311 (2020). (* equally contributed co-first authors)
- 4. Di Wang*, Huanyu Zhang*, Marco Gaboardi and Jinhui Xu. Estimating Smooth GLMs in Non-interactive Local Differential Privacy Model with Public Unlabeled Data. *The 32nd International Conference on Algorithmic Learning Theory (ALT 2021)*, Paris, France, March 16-19, 2021. (* equally contributed co-first authors)
- 5. Di Wang*, Hanshen Xiao*, Srinivas Devadas and Jinhui Xu. On Differentially Private Stochastic Convex Optimization with Heavy-tailed Data. The 37th International Conference on Machine Learning (ICML 2020), Vienna, Austria, July 12-18, 2020. (* equally contributed co-first authors)
- 6. Yunus Esencayi, Marco Gaboardi, Shi Li and **Di Wang**. Facility Location Problem in Differential Privacy Model Revisited. *Advances in Neural Information Processing Systems (NeurIPS 2019)*, Vancouver, BC, Canada, December 08-14, 2019. (Authors are alphabetically ordered)
- Di Wang and Jinhui Xu. On Sparse Linear Regression in the Local Differential Privacy Model. The 36th International Conference on Machine Learning (ICML 2019), Long Beach, CA, USA, June 9-15, 2019.
- 8. **Di Wang**, Changyou Chen and Jinhui Xu. Differentially Private Empirical Risk Minimization with Non-convex Loss Functions. *The 36th International Conference on Machine Learning (ICML 2019)*, Long Beach, CA, USA, June 9-15, 2019.
- Di Wang, Adam Smith and Jinhui Xu. Noninteractive Locally Private Learning of Linear Models via Polynomial Approximations. Algorithmic Learning Theory (ALT 2019), March 22-24, 2019, Chicago, IL, USA.
- 10. **Di Wang**, Marco Gaboardi and Jinhui Xu. Empirical Risk Minimization in Non-interactive Local Differential Privacy Revisited. *Advances in Neural Information Processing Systems (NeurIPS 2018)*, Montreal, QC, Canada, December 03-08, 2018.
- 11. **Di Wang**, Mingwei Ye and Jinhui Xu. Differentially Private Empirical Risk Minimization Revisited: Faster and More General. *Advances in Neural Information Processing Systems (NeurIPS 2017)*, Long Beach, CA, USA, 4-9 December 2017.

PUBLICATIONS

(* equally contributed co-first authors, '_____' students/postdocs/interns supervised by me)

Peer-Refereed Conference Papers

- 1. Di Wang*, Huanyu Zhang*, Marco Gaboardi and Jinhui Xu. Estimating Smooth GLMs in Non-interactive Local Differential Privacy Model with Public Unlabeled Data. *The 32nd International Conference on Algorithmic Learning Theory (ALT 2021)*, Online, March 16-19, 2021.
- 2. Mengdi Huai, Chenglin Miao, Jinduo Liu, Di Wang, Jingyuan Chou, and Aidong Zhang. Global Interpretation for Pairwise Learning. *The IEEE International Conference on Bioinformatics and Biomedicine 2020 (BIBM 2020)*, Online, December 16-19, 2020. (Selected as Regular Paper, Acceptance Rate: 19.4%).
- 3. Di Wang and Jinhui Xu. Escaping Saddle Points of Empirical Risk Privately and Scalably via DP-Trust Region Method. 2020 European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Database (ECML-PKDD 2020), Ghent, Belgium, September 14-18, 2020.
- 4. Di Wang*, Hanshen Xiao*, Srinivas Devadas and Jinhui Xu. On Differentially Private Stochastic Convex Optimization with Heavy-tailed Data. The 37th International Conference on Machine Learning (ICML 2020), Vienna, Austria, July 12-18, 2020.
- Mengdi Huai*, Di Wang*, Chenglin Miao, Jinhui Xu and Aidong Zhang. Pairwise Learning with Differential Privacy Guarantees. The Thirty-Fourth AAAI Conference on Artificial Intelligence (AAAI 2020), New York, USA, February 7-12, 2020.
- Di Wang*, Xiangyu Guo*, Chaowen Guan, Shi Li and Jinhui Xu. Scalable Estimating Stochastic Linear Combination of Non-linear Regressions. The Thirty-Fourth AAAI Conference on Artificial Intelligence (AAAI 2020), New York, USA, February 7-12, 2020.
- Mengdi Huai, Di Wang, Chenglin Miao and Aidong Zhang. Learning to Explain Pairwise Algorithms. The Thirty-Fourth AAAI Conference on Artificial Intelligence (AAAI 2020), New York, USA, February 7-12, 2020.
- 8. Yunus Esencayi, Marco Gaboardi, Shi Li and **Di Wang**. Facility Location Problem in Differential Privacy Model Revisited. *Advances in Neural Information Processing Systems (NeurIPS 2019)*, Vancouver, BC, Canada, December 08-14, 2019. (Authors are alphabetically ordered)
- 9. Mengdi Huai, Di Wang, Chenglin Miao, Jinhui Xu and Aidong Zhang. Privacy-aware Synthesizing for Crowdsourced Data. *The Twenty-Eighth International Joint Conference on Artificial Intelligence (IJCAI 2019)*, August 10-16, 2019, Macao, China.
- Di Wang and Jinhui Xu. Principal Component Analysis in the Local Differential Privacy Model. The Twenty-Eighth International Joint Conference on Artificial Intelligence (IJCAI 2019), August 10-16, 2019, Macao, China.
- 11. **Di Wang** and Jinhui Xu. Lower Bound of Locally Differentially Private Sparse Covariance Matrix Estimation. *The Twenty-Eighth International Joint Conference on Artificial Intelligence (IJCAI 2019)*, August 10-16, 2019, Macao, China.
- 12. **Di Wang** and Jinhui Xu. On Sparse Linear Regression in the Local Differential Privacy Model. *The 36th International Conference on Machine Learning (ICML 2019)*, Long Beach, CA, USA, June 9-15, 2019. (Selected as Long Talk, Acceptance Rate: 140/3424= 4.1%)
- 13. **Di Wang**, Changyou Chen and Jinhui Xu. Differentially Private Empirical Risk Minimization with Non-convex Loss Functions. *The 36th International Conference on Machine Learning (ICML 2019)*, Long Beach, CA, USA, June 9-15, 2019.
- Di Wang, Jinhui Xu and Yang He. Estimating Sparse Covariance Matrix Under Differential Privacy via Thresholding. The 53rd Annual Conference on Information Sciences and Systems (CISS 2019), Baltimore, MD, USA, March 20-22 2019.

- Di Wang, Adam Smith and Jinhui Xu. Noninteractive Locally Private Learning of Linear Models via Polynomial Approximations. Algorithmic Learning Theory (ALT 2019), March 22-24, 2019, Chicago, IL, USA.
- 16. **Di Wang** and Jinhui Xu. Differentially Private Empirical Risk Minimization with Smooth Non-Convex Loss Functions: A Non-Stationary View. *The Thirty-Third AAAI Conference on Artificial Intelligence (AAAI 2019)*, Honolulu, Hawaii, USA, January 27-February 1, 2019. (**Selected as Oral Presentation, Acceptance Rate: 460/7095=6.5%**)
- 17. **Di Wang**, Marco Gaboardi and Jinhui Xu. Empirical Risk Minimization in Non-interactive Local Differential Privacy Revisited. *Advances in Neural Information Processing Systems (NeurIPS 2018)*, Montreal, QC, Canada, December 03-08, 2018.
- 18. **Di Wang**, Mengdi Huai and Jinhui Xu. Differentially Private Sparse Inverse Covariance Estimation. 2018 IEEE Global Conference on Signal and Information Processing (GlobalSIP 2018), Anaheim, California, USA, November 26-29, 2018.
- Di Wang and Jinhui Xu. Large Scale Constrained Linear Regression Revisited: Faster Algorithms via Preconditioning. The Thirty-Second AAAI Conference on Artificial Intelligence (AAAI 2018), New Orleans, Louisiana, USA, February 2-7, 2018. (Selected as Oral Presentation, Acceptance Rate: 411/3800=10.8%)
- Di Wang, Mingwei Ye and Jinhui Xu. Differentially Private Empirical Risk Minimization Revisited: Faster and More General. Advances in Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA, 4-9 December 2017.

Peer-Refereed Journal Papers

- 21. **Di Wang** and Jinhui Xu. Differentially Private High Dimensional Sparse Covariance Matrix Estimation. *Theoretical Computer Science*, Volume 865, 14 April 2021, Pages 119-130.
- 22. **Di Wang** and Jinhui Xu. Inferring Ground Truth for Crowdsourcing Data Under Local Attribute Differential Privacy. *Theoretical Computer Science*, Volume 865, 14 April 2021, Pages 85-98.
- 23. **Di Wang** and Jinhui Xu. Sparse Linear Regression in the Local Model of Differential Privacy. *IEEE Transactions on Information Theory*, Volume 67, No. 2, Pages 1182-1200, Feb. 2021.
- 24. **Di Wang***, **Xiangyu Guo***, Shi Li and Jinhui Xu. Robust High Dimensional Expectation Maximization Algorithm via Trimmed Hard Thresholding. *Machine Learning* 109, 2283-2311 (2020).
- 25. **Di Wang**, Marco Gaboardi, Adam Smith and Jinhui Xu. Empirical Risk Minimization in the Non-interactive Local Model of Differential Privacy. *Journal of Machine Learning Research*, Volume 21, 200 (2020), Pages 1-39.
- Di Wang*, Xiangyu Guo*, Chaowen Guan, Shi Li and Jinhui Xu. Estimating Stochastic Linear Combination of Non-linear Regressions Efficiently and Scalably. *Neurocomputing*, Volume 399, 25 July 2020, Pages 129-140.
- 27. **Di Wang** and Jinhui Xu. Tight Lower Bound of Sparse Covariance Matrix Estimation in the Local Differential Privacy Model. *Theoretical Computer Science*, Volume 815, 2 May 2020, Pages 47-59.
- 28. **Di Wang** and Jinhui Xu. Principal Component Analysis in Local Differential Privacy Model. *Theoretical Computer Science*, Volume 809, 24 February 2020, Pages 296-312.
- 29. **Di Wang** and Jinhui Xu. Faster Constrained Linear Regression via Two-step Preconditioning. *Neurocomputing*, Volume 364, 28 October 2019, Pages 280-296.

Submitted Journal Papers

- 30. **Di Wang***, <u>Lijie Hu*</u>, <u>Huanyu Zhang*</u>, Marco Gaboardi and Jinhui Xu. Generalized Linear Model in Local Differential Privacy Model with Public Data. Submitted to *Journal of Machine Learning Research*.
- 31. **Di Wang***, **Hanshen Xiao***, Srinivas Devadas and Jinhui Xu. Differentially Private Heavy-tailed Stochastic Convex Optimization. Submitted to *IEEE Transactions on Information Theory*.
- 32. **Di Wang** and Jinhui Xu. Gradient Complexity and Non-Stationary Views of Differentially Private Empirical Risk Minimization. Submitted to *Computer and Security*.
- 33. **Di Wang***, **Mengdi Huai***, Chenglin Miao and Jinhui Xu. Differentially Private Pairwise Learning. Submitted to at *Neural Networks*.

Submitted Conference Papers

- 34. Hui Zhang, Peng Zhao and Di Wang. Differentially Private Stochastic Convex Optimization Revisit.
- 35. Youming Tao, Yulian Wu and **Di Wang**. Optimal Differentially Private Heavy-tailed Multi-Arm Bandits.
- 36. <u>Lijie Hu, Shuo Ni</u> and **Di Wang**. High Dimensional Differentially Private Heavy-tailed Stochastic Convex Optimization.

Workshop Papers

- 37. Di Wang*, Huanyu Zhang*, Marco Gaboardi and Jinhui Xu. Estimating Smooth GLMs in Non-interactive Local Differential Privacy Model with Public Unlabeled Data. NeurIPS Workshop on Privacy in Machine Learning, 2019.
- 38. **Di Wang**, Adam Smith and Jinhui Xu. High Dimensional Sparse Linear Regression under Local Differential Privacy: Power and Limitations. *NeurIPS Workshop on Privacy Preserving Machine Learning*, 2018.

Students

- 1. Lijie Hu (CS PhD), Invariant Learning, Fairness, 01/2021-
- 2. Zihang Xiang (CS PhD), Private Learning System, Machine Unlearning 01/2021-
- 3. Yulian Wu (CS PhD), Differential Privacy, Bandit, Reinforcement Learning 09/2021-
- 4. Chenglong Wang (CS PhD), Machine Unlearning, 09/2021-

Visiting Students/Research Intern

- 1. Zejun Xie (Undergraduate at Renmin University of China), 07/2020-08/2020.
- Zhiyu Xue (Undergraduate at University of Electronic Science and Technology of China), 08/2020-10/2020.
- 3. Shaoyang Yang (Undergraduate at Harbin Institute of Technology), 08/2020-10/2020.
- 4. Youming Tao (Undergraduate at Shandong University), 01/2021-06/2021.
- 5. Shuo Ni (Master student at University of South California), 01/2021-06/2021
- 6. Xingyu Jiang (Undergraduate at Harbin Institute of Technology Weihai), 01/2021-06/2021
- Hui Zhang (Undergraduate at University of Electronic Science and Technology of China), 03/2021-
- 8. Mingyi Zhou (Master student at University of Electronic Science and Technology of China), 04/2021-

TALKS

INVITED TALKS

- How To Preserve Privacy In Learning?
 - 1. School of Cyber Science and Technology, Zhejiang University, September 2020
 - 2. School of Computing and Information Systems, University of Melbourne, April 2020
 - 3. Department of Computer Science and Engineering, Chinese University of Hong Kong, April 2020
 - 4. Department of Computer Science, Dalhousie University, April 2020
 - 5. CISPA-Helmholtz Center for Information Security, April 2020
 - 6. Department of Computing, Hong Kong Polytechnic University, April 2020
 - 7. Department of Computer Science, University of Memphis, Apirl 2020
 - 8. School of Computer Science, University of Sydney, April 2020
 - 9. Department of Computing, Imperial College London, March 2020
 - 10. King Abdullah University of Science and Technology, March 2020
 - 11. Department of Computing and Software, McMaster University, March 2020
 - 12. Department of Computer Science, City University of Hong Kong, March 2020
 - 13. School of Information System, Singapore Management University, March 2020
 - 14. Department of Computer Science and Engineering, Hong Kong University of Science and Technology, February 2020
 - 15. Department of Computer Science, McGill University, February 2020
 - 16. Department of Computer Science, University of Alberta, November 2019
- Differentially Private Machine Learning
 - 17. Department of Computer Science, University College London, UK, March 2020
 - 18. Department of Computer Science, University of Warwick, UK, March 2020
 - 19. School of Computer Science, University of Birmingham, UK, March 2020
 - 20. Department of Computer Science, University of Surrey, UK, February 2020
- Fundamental Machine Learning Problems in Differential Privacy Model
 - 21. Department of Computer Science, University of Science and Technology of China, November 2019
 - 22. Department of Computer Science, Nanjing University, China, November 2019

CONFERENCE TALKS

- Estimating Smooth GLMs in Non-interactive Local Differential Privacy Model with Public Unlabeled Data. ALT 2021, Online.
- 2. Robust High Dimensional Expectation Maximization Algorithm via Trimmed Hard Thresholding. ACML 2020, Online.
- 3. Escaping Saddle Points of Empirical Risk Privately and Scalably via DP-Trust Region Method. ECML-PKDD 2020, Online.
- 4. On the Differentially Private Stochastic Optimization with Heavy-tailed Data. ICML 2020, Online.
- Principal Component Analysis in the Local Differential Privacy Model. IJCAI 2019. Macau, China, August, China (Online).

- 6. Lower Bound of Locally Differentially Private Sparse Covariance Matrix Estimation. IJCAI 2019. Macau, China, August, China (Online).
- 7. On the Locally Differentially Private Sparse Linear Regression. ICML 2019. Long Beach, CA, USA. June 2019.
- 8. Estimation Sparse Covariance Matrix Under Differential Privacy via Thresholding. CISS 2019. Baltimore, MD, USA. March 2019.
- 9. Empirical Risk Minimization in Non-interactive Local Model via Polynomial of Inner Product Approximation. ALT 2019. Chicago, IL, USA. March 2019.
- 10. Differentially Private Sparse Inverse Covariance Estimation. 2018 IEEE GlobalSIP Signal Processing for Adversarial Machine Learning. November, 2018.
- 11. Differentially Private Empirical Risk Minimization in the Non-interactive Local Model, Intern Presentation, Harvard University, June, 2018.
- 12. Large Scale Constrained Linear Regression Revisited Faster Algorithms via Preconditioning, The Thirty-Second Conference on Artificial Intelligence (AAAI), February, 2018.
- 13. Differentially Private Empirical Risk Minimization with Non-convex Loss Function, SUNY Buffalo CSE 50th Anniversary, University at Buffalo, September, 2017.

PROFESSIONAL SERVICE

- Technical Program Committee Member:
 - The 26th European Symposium on Research in Computer Security (ESORICS), 2021 (Poster Session)
 - AAAI Conference on Artificial Intelligence (AAAI), 2021
 - Winter Conference on Applications of Computer Vision (WACV 2020)
 - European Conference on Machine Learning (ECML-PKDD 2020)
 - The 29th International Joint Conference on Artificial Intelligence (IJCAI-PRICAI 2020)
 - IEEE Symposium on Security and Privacy 2020 (Shadow PC)
 - AAAI Conference on Artificial Intelligence (AAAI), 2020
- Reviewer (Journals)
 - Information Science
 - Neurocomputing
 - IEEE Transactions on Big Data
 - ACM Computing Surveys
 - IEEE Transactions on Information Forensics and Security
 - IEEE Transactions on Pattern Analysis and Machine Intelligence
 - Theoretical Computer Science
 - Information Processing Letters
 - Security and Communication Networks
 - Patterns
- Reviewer (Conferences)
 - Neural Information Processing Systems (NeurIPS/NIPS) 2021
 - 2021 IEEE International Symposium on Information Theory (ISIT 2021)
 - IEEE International Conference on Computer Vision (ICCV 2021)

- The 38th International Conference on Machine Learning (ICML 2021)
- The 24th International Conference on Artificial Intelligence and Statistics (AISTATS 2021)
- IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2021)
- 2021 International Conference on Learning Representations (ICLR 2021)
- The 15th Asian Conference on Computer Vision (ACCV 2020)
- Neural Information Processing Systems (NeurIPS/NIPS) 2020
- The 16th Annual Conference on Theory and Applications of Models of Computation (TAMC 2020)
- The 36th International Symposium on Computational Geometry (SoCG 2020)
- European Conference on Computer Vision (ECCV 2020)
- The 52nd ACM Symposium on Theory of Computing (STOC 2020)
- IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2020)
- Neural Information Processing Systems (NeurIPS/NIPS) 2019
- IEEE International Conference on Distributed Computing Systems (ICDCS 2019)
- IEEE International Conference on Computer Vision (ICCV 2019)
- IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2019)
- International Conference on Machine Learning (ICML) 2019
- International Conference on Artificial Intelligence and Statistics (AISTATS) 2019
- AAAI Conference on Artificial Intelligence (AAAI) 2018
- ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD) 2018
- AAAI Conference on Artificial Intelligence (AAAI) 2017
- International Symposium CompIMAGE'18-Computational Modeling of Objects Presented in Images: Fundamentals, Methods, and Applications
- International Workshop on Combinatorial Image Analysis (IWCIA) 2017

ACADEMIC THESIS COMMITEE

KAUST SERVICE

- 2021 KAUST Gifted Student Programs Convocation, February 2021
 - Sci Café: Dynamic presentation to showcase an area of KAUSTs innovative research, AI & Cyber Security.
 - Faculty Mentoring Meetings: Meet one-on-one with junior and senior KGSP students to provide guidance and feedback on the students professional and academic development, areas of strengths and weakness, and recommendations for future activities.